

الفصل الثاني

المجتمع والعينات

The Population and Samples

2-1: مقدمة:

الهدف الرئيسي للاحتمال هو الحصول على استقرار أو تنبؤ عن المجتمع محل الدراسة من واقع معلومات محتواة في عينة من هذا المجتمع تستخدم لتقدير معالم المجتمع مثل الوسط الحسابي ، المجموع ، و النسبة ، و عدد أفراد ، و التباين وغيرها .

سنبدأ موضوع المعاينة بتعريف عناصرها، فمن الملاحظ أن كل مشاهدة تحتوي على كمية من المعلومات عن الصفة المدروسة للمجتمع الذي أخذت منه العينة ولأن المعلومات لها تكلفة فإن الباحث يجب أن يحدد حجم العينة اللازمة لكمية المعلومات التي يحتاج إليها ضمن الميزانية المقررة . ومن المؤكد أن قلة المعلومات لا تعطي تقديراً جيداً للمعلمة أو للصفة المدروسة ، وقد تكون كثرة المعلومات تبذيراً للمخصصات المتوفرة .

إن نوعية المعلومات التي نحصل عليها من العينة تعتمد على حجمها (عدد عناصرها) وعلى تباعدها بعضها عن بعض ؛ إذ يمكن التحكم بتباينها عن طريق اختيار العينة ، وهذا ما يسمى بتصميم التجربة. ويمكن أن نشير إلى أسلوبين لجمع المعلومات هما أسلوب الحصر الشامل وأسلوب العينة.

2-1-1: الحصر الشامل والعينة:

أولاً: الحصر الشامل (Complete Enumeration) :

هذه الطريقة تعني دراسة شاملة لجميع أفراد المجتمع محل الدراسة ، ولكن هذه الطريقة عالية التكلفة، وتحتاج إلى جهود وإمكانيات ضخمة، وهذا النوع من الدراسات يتم عادة على فترات متباعدة مثل تعدادات السكان والتعدادات الزراعية.... تستخدم بيانات الحصر الشامل في التخطيط لمختلف البرامج السكانية والصحية للدولة.

ثانياً: العينة (The sample) :

العينة هي مجموعة من المفردات يتم اختيارها من المجتمع محل الدراسة بشكل عشوائي، إذ إنّ دراسة المجتمع بطريقة الحصر الشامل تبدو صعبة في كثير من الأحيان على الرغم من دقتها في أغلب الأحوال ، فإن معظم الدراسات تتم بطريقة العينة بدلاً من الحصر الشامل.

لكي نحكم على الكل (أي المجتمع) باستخدام الجزء حكماً دقيقاً يجب أن نهتم بطرق اختيار هذا الجزء الذي يسمى بالعينة (Sample). وتسمى عملية اختيار هذه العينة بالمعاينة (Sampling).

من الضروري اختيار العينة بحيث تكون ممثلة للمجتمع أي تتصف بنفس صفاته، وذلك لنستطيع تعميم نتائجها على المجتمع. من خلال دراسة تلك العينة يمكن الاستدلال على خصائص المجتمع الذي أخذت منه عن طريق تقدير معالمه (Parameter) ونظرية العينات تربط بين المجتمع والعينة حيث يمكن تقدير متوسط أو مجموع أو نسبة ما للمجتمع أو فترات الثقة لهذه المعالم باستخدام

مقاييس مأخوذة من العينة تسمى الإحصاءات (Statistics).

2-1-2: أسباب تفضيل العينة على الحصر الشامل:

المعينة هي علم وفن التحكم وقياس دقة المعلومات الإحصائية باستخدام النظريات العلمية ، وأصبح من المألوف الاعتماد على العينات في الدراسات المختلفة ، وقد يعتقد أن استخدام العينات يقلل من دقة المعلومات ، لكن اختيار العينات بشكل صحيح يؤدي إلى نتائج دقيقة (قريبة من الواقع) قد لا تقل عن أسلوب الحصر الشامل، لأسباب عديدة أهمها:

- 1) محدودية الإمكانيات لإجراء الحصر الشامل.
- 2) طريقة العينات تقلل التكلفة والجهد والزمن.
- 3) صعوبة حصر أفراد المجتمع.
- 4) تلف مفردات المجتمع المدروسة.
- 5) في المجتمع المتجانس لا مسوغ لدراسته بطريقة الحصر.

مع كل هذه النقاط الإيجابية لأسلوب العينة عليها مآخذ نذكر منها:

- 1) مهما بلغت الدقة في استخدام العينات تَبْقَى النتائج تقديرية.
- 2) استخدام العينات يحتاج إلى كوادر فنية مدربة ومؤهلة.
- 3) استخدام العينات يحتاج إلى تخطيط وإعداد وتنفيذ ، ومن ثم تحليل المعلومات بعد الحصول عليها.

2-2: الخطوات الرئيسية في تصميم العينات (Basic Steps for sampling design):

إن اختيار عينة من 1000 شخص مرتبين في ملف أمر سهل جداً، لكن اختيار العينة من منطقة أو عدة مناطق متفرقة تحتاج إلى التنقل بينها بالإضافة إلى غياب رغبة الأشخاص في إعطاء المعلومات ، وهو أمر أصعب بكثير.

لنعرض الآن الخطوات الرئيسية لاستخدام العينات:

أولاً : هدف الدراسة (Objective):

تحديد الهدف من الدراسة من أهم الخطوات ، لأنها تحدد المطلوب، ولاسيما إذا كانت الدراسة معقدة ، فقد تتسبب التفاصيل الهدف الرئيسي.

ثانياً : تحديد المجتمع محل المعاينة (The population):

يقصد بالمجتمع الأفراد أو العناصر أو المفردات أو الوحدات التي نرغب بدراستها، فقد يتكون المجتمع من أشخاص . مرضى . منتجات طبية . مشافٍ ... إلخ. فإذا كنا نرغب في تقدير كمية المادة الفعالة في أحد أنواع المسكنات من إنتاج شركة أدوية محددة فإن المجتمع المدروس معروف ومحدد تماماً . أما إذا كنا نرغب في تقدير متوسط وزن الأطفال عند الولادة ، فإن المجتمع الذي سنختار منه العينة غير محدد : هل نريد الذكور فقط أم الإناث ؟ أم مواليد هذه المنطقة أو تلك المدينة ؟... ويجب أن نتذكر أن المفردات تكون المجتمع ووحدة المعاينة (Sampling Unit) هي عبارة عن عنصر من العينة وقد تكون شخصاً - أو نباتاً معيناً - أو مشفىً - أو أسرةً - أو مدينةً.

ثالثاً : الإطار (The Frame) :

قبل اختيار العينة يجب تقسيم المجتمع إلى وحدات معاينة، وهذه الوحدات يجب ألا تكون متداخلة. تسمى القائمة التي تحتوي على هذه الوحدات بالإطار. وتحديد الإطار خطوة مهمة في المعاينة ، وقد يكون الإطار قائمة بأفراد مجتمع . خريطة منطقة . صناديق من منتج معين . قائمة بأسماء شركات الأدوية في بلد . قائمة بأسماء المشافي الخاصة . قائمة بأسماء أطباء الأسنان .

رابعاً : جمع البيانات (Data collection) :

يجب أن تتوافق البيانات التي تجمع مع هدف الدراسة ، وأن لا تحتوي الاستمارة على تفاصيل كثيرة، الأمر الذي يقلل من جودة الإجابات.

خامساً : حجم العينة (Sample Size) :

- (a) يختلف حجم العينة من مجتمع لآخر ، فإذا كان المجتمع متجانساً فلا بأس أن يكون حجم العينة صغيراً، مثال ذلك فحص الدم. فنقاط قليلة كافية لإجراء التحليل. أما إذا لم يكن متجانساً فيجب زيادة حجم العينة . فإذا كنا نرغب في تقدير معدل الأمطار في الجمهورية العربية السورية يجب أخذ متوسط عدد كبير من المناطق ، وذلك بسبب تفاوت المعدلات في مختلف المناطق.
- (b) إنَّ دقة الدراسة وصحة النتائج متناسبة مع حجم العينة ، فكلما كان حجم العينة أكبر كانت النتائج أقرب إلى الواقع.
- (c) الميزانية المقررة للدراسة تحدد حجم العينة. إذا كانت المخصصات قليلة فلا بد من تصغير حجم العينة ، وإذا كانت كبيرة فالأفضل زيادة حجم العينة.

سادساً : طرق جمع البيانات (Methods of data collection) :

لإجراء أي دراسة لا بد من جمع بيانات عن الظاهرة المدروسة بعد تحديد هدف الدراسة، وهناك مصدران تجمع منهما البيانات ، مصدر داخلي وآخر خارجي (الفصل الثالث).

2-3:أنواع العينات وطرق المعاينة (Sample Type and sampling techniques) :

إذا كان الهدف مراقبة التغيرات التي تطرأ على المتغيرات خلال فترات زمنية مثل جودة دواء ما . معدلات الإصابة بأحد الحماح الراشحة فإن العينة المختارة تظل ثابتة (تحتوي على نفس وحدات المعاينة) على مدار الفترة الزمنية للدراسة. أما النوع الشائع للعينات والأكثر استخداماً هو العينات المتغيرة (التي لا تحوي على نفس وحدات المعاينة إلا بالصدفة) فأخذ عينات ثابتة يؤدي إلى الضجر والملل، الأمر الذي قد ينعكس سلباً على نوعية وجودة البيانات. لذا ينصح بتجديد العينة أو تجديد جزء منها على الأقل.

للمعاينة طرق متعددة تعتمد على نوعية المجتمع المراد دراسته وعلى الهدف من الدراسة، وتنقسم العينات إلى عينات عشوائية وعينات غير عشوائية.

2-3-1: العينات العشوائية (Random Sample) :

يتم في هذا النوع من العينات اختيار أفراد العينة من المجتمع بطريقة غير متحيزة بحيث نضمن لكل فرد من المجتمع نفس الإمكانية في الظهور في العينة، وهذا يضمن إمكانية إخضاع هذا النوع من العينات للقوانين الاحتمالية ومن أنواع

العينات العشوائية:

1) العينات العشوائية البسيطة (Simple Random Sample) واستخدام جداول الأرقام العشوائية:

العينات العشوائية البسيطة أبسط أنواع العينات ، وتستخدم في حالة تجانس أفراد المجتمع محل الدراسة في الظاهرة المدروسة ومعرفة جميع أفرادها. فيتم اختيار الأفراد إما بإجراء قرعة ، إذ يتم ترقيمهم ثم اختيار الأفراد بسحب أرقام الأفراد الداخلين في العينة بطريقة غير متحيزة وذلك بكتابة الأرقام على أوراق متشابهة وخلطها جيداً ثم سحب العدد المطلوب. وهناك طريقة أخرى أفضل وهي استخدام جداول الأرقام العشوائية ولاسيماً حين يكون حجم العينة كبيراً.

يتكون جدول الأرقام العشوائية من مجموعة من الأعداد المكونة من خمس خانات مرتبة في صفوف وأعمدة ، وتحتوي على مجموعة الأرقام من 0 إلى 9 بنسب متساوية ، وإن اختيار رقم من هذا الجدول يكافئ سحب ورقة بشكل عشوائي من مجموعة الأوراق المخلوطة جيداً التي تحمل الأرقام من 0 إلى 9، ونوضح في المثال التالي طريقة استخدام هذا الجدول.

مثال (1-2):

نريد اختيار عينة عشوائية بسيطة مكونة من 10 مرضى من الذين زاروا إحدى العيادات الخارجية في مشفى الأسد الجامعي ، وذلك خلال العام الماضي ، فإذا افترضنا أن عددهم 600 ، يحمل كل اسم رقماً من 1 إلى 600.

نختار بشكل عشوائي صفحة من الجدول ومن تلك الصفحة نحدد بشكل عشوائي سطراً وعموداً ، ونبدأ من العدد الواقع في تقاطع السطر والعمود المختارين ،

ونختار منه أول ثلاث خانات من اليسار ؛ (لأن العدد 600 يتكون من ثلاث خانات) ، ثم ننقل إلى عدد ثانٍ وثالث بترتيب نحدده مسبقاً سواء إلى اليمين أم اليسار أم الأعلى أم الأسفل ، ونسجل قائمة بالأعداد التي تم اختيارها ، ثم نحذف كل عدد يزيد على 600.

لو كان رقم السطر المختار 12 والعمود 5 وجدنا العدد 02338 يقع في السطر 12 والعمود 5 . نأخذ منه أول ثلاثة أرقام 338 ولو حددنا جهة الانتقال مثلاً إلى اليسار ، وتابعنا في قراءة الأعداد نسجل (338 ، 772 ، 774 ، 165 ، 931 ، 812 ، 153 ، 090 ، 649 ، 754 ، 822 ، 924 ، 515 ، 917 ، 655 ، 174 ، 927 ، 458 ، 590 ، 393 ، 286). نهمل الأعداد التي تزيد على 600 ، فتكون العينة مكونة من المرضى الذين أرقامهم:

(338 ، 165 ، 153 ، 090 ، 515 ، 174 ، 458 ، 590 ، 393 ، 286).

(2) العينة العشوائية الطبقية (Stratified Random Sample):

عندما يكون المجتمع غير متجانس نستخدم هذا النوع من العينات ؛ إذ نقسم المجتمع إلى مجموعات متجانسة تسمى طبقات ، ونختار عينة عشوائية بسيطة من كل طبقة. فإذا أردنا دراسة متوسط إنفاق الطلاب على المكالمات الهاتفية إذ نتوقع أن معدلات الإنفاق متجانسة ضمن كل كلية ، وتختلف بين كلية وأخرى، لذلك نقسم الطلبة إلى طبقات حسب الكليات ، ونختار عشوائياً من كل كلية (طبقة) عينة عددها يتناسب مع عدد طلبتها نسبة إلى عدد طلبة الجامعة ، وندرس متوسط الإنفاق للطلاب الذين تم اختيارهم من أجل ضمان تمثيل الطبقات الصغيرة.

(3) العينة العشوائية المنتظمة (Systematic Random Sample):

تستخدم هذه العينة عندما يكون المجتمع متجانساً ومرتباً وفقاً لصفة معينة، ونريد أن نختار فرداً من كل عدد محدد من الأفراد (وليكن 10 مثلاً). فنختار الفرد الأول عشوائياً ، فإذا افترضنا أن رقمه 7 نختار الفرد الثاني ذا الرقم 17 والثالث ذا الرقم 27 وهكذا حتى يكتمل العدد المطلوب للعينة ، ويعتمد اختيار كل فرد على حجم العينة وحجم المجتمع.

(4) العينة العشوائية العنقودية (Cluster Random Sample):

يقسم المجتمع إلى مجموعات صغيرة تكون متجانسة ، ثم نختار عينة بشكل عشوائي من هذه المجموعات ، وتكون أفراد المجموعات المختارة هي عناصر العينة العنقودية ، وندعوها بالعينة العنقودية العشوائية من مرحلة واحدة. أما إذا اخترنا عينة عشوائية بسيطة من كل مجموعة من المجموعات المختارة ، فتسمى بالعينة العنقودية العشوائية من مرحلتين. مثال على ذلك إذا أردنا اختيار عينة عشوائية لنوع من الدواء الذي تنتجه شركه ما. وإذا كانت الشركة تنتج يومياً خمسة صناديق، يحوي الصندوق الواحد 100 علبة نختار بشكل عشوائي مجموعة من الصناديق ، وليكن عددها 10 صناديق ، ثم نختار من كل صندوق 10 علب لنحصل على عينة من مئة علبة نجري عليها الفحص اللازم.

2-3-2: العينات غير العشوائية (Nonrandom Sample):

يتم اختيار العينة في هذا النوع من العينات بشكل متعمد ، ونجري الدراسة على مجموعة محددة لا يخضع اختيارها للقرعة. من هذه العينات:

(1) العينة المصادفة (Accidental Sample):

إذا أجرى طبيب دراسة على مجموعة من المرضى الذين يراجعونه تكون هذه المجموعة عينة اختارتها الصدفة ، وليس للباحث أثر في اختيارها .

(2) العينة الحصصية (Quota Sample):

يقوم الباحث بتقسيم المجتمع إلى مجموعات ، ثم ينتقي من كل مجموعة فئة صغيرة ممثلة له يختارها حسب معيار ما وليس عشوائياً .

(3) العينة القصدية (العمدية) (Purposive Sample):

يختار الباحث مجموعة من الأفراد حسب ما يراه مناسباً لتحقيق هدف معين. فقد يختار أحد الأطباء الباحثين مجموعة محددة من الأطباء يختارهم كيفما يريد لمعرفة رأيهم بأحد أنواع العقاقير .

2-4: تقدير وسطاء (معالم) المجتمع في العينات العشوائية البسيطة:

2-4-1: تقدير المتوسط:

بفرض أن حجم المجتمع N وأن μ ، σ^2 هما متوسط المجتمع وتباينه ، فإذا كان حجم العينة n نرمز بـ \bar{x} و S^2 لمتوسط وتباين العينة فكما نعلم أن $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ هو مقدر لمتوسط المجتمع μ .

وكذلك نستخدم العلاقة الآتية لتقدير تباين المتوسط \bar{x} :

$$\hat{v}(\bar{x}) = \frac{S^2}{n} \left(\frac{N-n}{N} \right) \quad (2-1)$$

وتقدير الخطأ المعياري (Standard Error) هو :

$$\sqrt{\hat{v}(\bar{x})} = \sqrt{\frac{S^2}{n} \left(\frac{N-n}{N} \right)} \quad (2-2)$$

ويكون حد الخطأ في التقدير :

$$\beta = 2\sqrt{\hat{v}(\bar{x})} = 2\sqrt{\frac{S^2}{n} \left(\frac{N-n}{N} \right)} \quad (2-3)$$

ويسمى $\frac{N-n}{N} = 1 - \frac{n}{N}$ بمعامل التصحيح للمجتمع ، ويساوي $1 - f$ حيث f نسبة العينة إلى المجتمع. فإذا كانت $f \geq 0.95$ أو $0.95 N \leq n$ أي حجم العينة لا يقل عن 95% من حجم المجتمع نعتبر معامل التصحيح يساوي 1 .
ويصبح الخطأ المعياري $\sqrt{\frac{S^2}{n}}$ ويحدد لنا حد الخطأ الذي يصبح $2\sqrt{\frac{S^2}{n}}$.
95% فترة ثقة لمتوسط المجتمع.

$$(\bar{x} - \beta \quad \bar{x} + \beta) \quad (2-4)$$

وبشكل عام لو رمزنا بـ α لعدد صغير قد يكون 0.01 ، 0.02 ، فإن المجال التالي هو $100\% (1 - \alpha)$ مجال ثقة لـ μ متوسط مجتمع طبيعي ومن أجل عينة كبيرة:

$$\bar{x} - z_{1-\alpha/2} \sqrt{\frac{S^2}{n} (1 - f)} < \mu < \bar{x} + z_{1-\alpha/2} \sqrt{\frac{S^2}{n} (1 - f)} \quad (2-5)$$

مثال (2-2):

أخذنا عينة عشوائية مكونة من 25 شخصاً من أحد بيوت المسنين لتقدير متوسط عدد ضربات القلب في الدقيقة. فكان متوسط العينة $\bar{x} = 70$ وانحرافها المعياري $S = 5$. فإذا علمنا عدد النزلاء يساوي $N = 150$. فقَدِّرْ متوسط ضربات القلب، ثم احسب حد الخطأ في التقدير، وأوجد 95% فترة ثقة للمتوسط μ .

الحل :

لدينا $N = 150$ ، $n = 25$ ، $\bar{x} = 70$ ، $S = 5$.

$\bar{x} = 70$ هو تقدير للمتوسط الحقيقي μ ، أما تقدير التباين حسب (1 - 5)

$$\hat{v}(\bar{x}) = \frac{25}{25} \left(\frac{150-25}{150} \right) = 0.83$$

أما حد الخطأ في التقدير: $\beta = 2\sqrt{\hat{v}(\bar{x})} = 2\sqrt{0.83} = 1.8$

حسب (4 - 2) نعتبر الفترة التي طرفاها $\bar{x} \pm 2\sqrt{\hat{v}(\bar{x})}$ فترة ثقة لمتوسط عدد ضربات القلب أي المتوسط الحقيقي لعدد ضربات القلب $68.2 < \mu < 71.8$ وذلك بثقة مقدارها 95%.

2-4-2: تقدير المجموع الكلي للمجتمع τ Estimation of the (population Total):

يمكن تقدير المجموع الكلي $\tau = \sum_{i=1}^N x_i$ بناءً على تقدير متوسط المجتمع \bar{x}

$$\hat{\tau} = N \cdot \bar{x} \quad \text{بالعلاقة:}$$

وتقدير تباين المجموع:

$$\hat{v}(\hat{t}) = N^2 \hat{v}(\bar{x}) = N^2 \frac{S^2}{n} \left(\frac{N-n}{N} \right) = \frac{S^2}{n} \cdot N(N-n) \quad (2-6)$$

وحد خطأ التقدير :

$$\beta = 2\sqrt{\hat{v}(\hat{t})} = 2N\sqrt{\hat{v}(\bar{x})} \quad (2-7)$$

مثال (2-3):

عينة عشوائية من 20 مريضاً من المرضى الذين دخلوا أحد المشافي الخاصة و الذين بلغ عددهم العام الماضي 1000 حسبنا متوسط وتباين قيم الفواتير التي سددها للمشفى، فكان $\bar{x} = 20000$ ، $S^2 = (1500)^2$. قدر مجموع مدفوعات جميع النزلاء في العام الماضي ، واحسب خطأ التقدير ، ثم أوجد 95% فترة ثقة لمجموع المدفوعات.

الحل:

تقدير المجموع الكلي: حجم العينة $n = 20$ وحجم المجتمع $N = 1000$

$$\hat{t} = 1000(20000) = 20000000 \text{ ليرة}$$

تقدير التباين من (2 - 6) :

$$\hat{v}(\hat{t}) = \frac{2250000}{20} (1000)(1000 - 20) = 11025000000 \text{ ليرة}$$

وحد الخطأ في التقدير :

$$\beta = 2\sqrt{\hat{v}(\hat{t})} = 664078 \text{ ليرة}$$

وتكون 95% فترة ثقة للمجموع τ مقدراً بالليرات

$$(\hat{\tau} - 2\beta < \tau < \hat{\tau} + 2\beta) =$$

$$(19335922 < \tau < 20664078)$$

أي نحن على ثقة مقدارها 95% من أن مجموع فواتير المرضى لا يقل عن 19.3 ولا يزيد على 20.6 مليون ليرة تقريباً.

3-4-2: تقدير نسبة النجاح p للمجتمع (Estimation of the) :(population proportion)

نعتبر المجتمع مكوناً من فئتين الأولى حجمها N_1 تحمل صفة ما والثانية حجمها N_2 لا تحمل تلك الصفة حيث : $N = N_1 + N_2$ مثلاً فئة الذكور وفئة الإناث أو المدخنين وغير المدخنين المصابين بمرض السكري والأصحاء.... إلخ .

فإذا رمزنا بـ P للنسبة الحقيقية للذين يحملون الصفة المدروسة $P = \frac{N_1}{N}$ و \hat{P} لنسبة الذين يحملون تلك الصفة في العينة $\hat{P} = \frac{n_1}{n}$ حيث $n = n_1 + n_2$.

n_1 : عدد الذين يحملون تلك الصفة في العينة . وبفرض $\hat{q} = 1 - \hat{P}$.
يكون تقدير تباين النسبة

$$\hat{v}(\hat{P}) = \frac{\hat{P} \cdot \hat{q}}{n-1} \left(\frac{N-n}{N} \right) \quad (2-8)$$

$$\beta = 2\sqrt{\hat{v}(\hat{P})} \quad \text{وحد خطأ التقدير فيساوي}$$

مثال (1-2):

بفرض أن عدد طلاب كلية الطب يساوي $N = 1500$ أخذنا منهم عينة حجمها $n = 100$ طالب وكان عدد المدخنين منهم 15 طالباً. قدر نسبة المدخنين في كلية الطب ، واحسب 95% مجال ثقة لنسبة المدخنين.

الحل:

$$\hat{P} = \frac{15}{100} = 0.15 \text{ إن مقدر نسبة المدخنين}$$

وتقدير تباين \hat{P} من (5 - 8) :

$$\hat{v}(\hat{P}) = \frac{(0.15)(0.85)}{99} \left(\frac{1500-100}{1500} \right) \cong 0.0012$$

وحد الخطأ للتقدير :

$$\beta = 2\sqrt{0.0012} \cong 0.07$$

و 95% فترة ثقة ل \hat{P} هي (0.08 0.22)

أي نحن على ثقة مقدارها 95% أن النسبة الحقيقية للمدخنين لا تقل عن 8% ولا تزيد على 22% .

2-4-4: اختيار حجم العينة n (Selection of the Sample size):

إن حجم العينة n له أهمية كبيرة في النتائج النهائية للدراسة ، فدائماً نريد تقديرات جيدة للوسطاء ، لذلك نحتاج إلى عينات حجمها كبير، لكننا في نفس الوقت لا نريد أن تكون النفقات عالية مقابل فائدة قليلة. فإذا أردنا تحديد حجم العينة n يجب أن نعلم قيم β ، σ^2 ، N حيث من خلال β نضع شرطاً على حد الخطأ

المسموح به. أما σ^2 فتكون غالباً مجهولة ، ولا بد من تقديرها إما من دراسات سابقة وإما من عينة أولية ولنقُم بتقدير حجم العينة n في التقديرات السابقة.

(1) حجم العينة اللازم لتقدير μ نعلم أن :

$$\beta = 2\sqrt{\hat{v}(\bar{x})} = 2\sqrt{\frac{\sigma^2}{n} \left(\frac{N-n}{N-1} \right)}$$

مقارنة $V(\bar{x})$ مع قيمته في العلاقة (2-2)

$$\text{بالتربيع نجد} \quad \beta^2 = 4 \cdot \frac{\sigma^2}{n} \left(\frac{N-n}{N-1} \right) \quad . \text{ومنها نعين } n \text{ بدلالة } \beta$$

$$n = \frac{N \cdot \sigma^2}{(N-1)D + \sigma^2} ; \quad D = \left(\frac{\beta}{2} \right)^2 \quad (2-9)$$

(2) حجم العينة اللازم لتقدير t مجموع قيم المجتمع

$$n = \frac{N \cdot \sigma^2}{(N-1)D + \sigma^2} ; \quad D = \left(\frac{\beta}{2N} \right)^2 \quad (2-10)$$

(3) حجم العينة اللازم لتقدير P

$$n = \frac{N \cdot P \cdot q}{(N-1)D + P \cdot q} ; \quad D = \left(\frac{\beta}{2} \right)^2 \quad (2-11)$$

إذا لم يتوفر تقدير لـ σ^2 نعتبر $\sigma = \frac{R}{4}$ ربع مدى البيانات ونعتبر $\hat{P} = \frac{1}{2}$ إذا كان مجهولاً.

مثال (5-2):

نريد دراسة متوسط أوزان المواليد الذكور في أحد المشافي باستخدام العينة العشوائية البسيطة. فإذا علمنا أن عدد المواليد الذكور في العام السابق بلغ 500 مولود ، وأن الانحراف المعياري تم تقديره بـ $\sigma = 100 \text{ g}$. قدر حجم العينة اللازم لهذه الدراسة إذا أردنا ألا يزيد حد خطأ التقدير على 10 غرامات؟

الحل:

من (9 - 2) لدينا

$$\beta = 10 , \sigma^2 = 10000 , N = 500$$

$$D = \frac{\beta^2}{4} = \frac{100}{4} = 25$$

$$n = \frac{N \cdot \sigma^2}{(N-1)D + \sigma^2} = \frac{500(10000)}{(499)(25) + 10000} \cong 223$$

إذن نحتاج إلى عينة حجمها 223 حتى لا يتجاوز حد الخطأ في تقدير المتوسط 10μ غرامات.

تمارين ومسائل

1. حدد جميع العينات المختلفة الممكنة التي حجمها $n = 3$ والتي يمكن اختيارها من المجتمع المكون من الأرقام $\{2, 4, 6, 8, 10\}$ ثم:

أ- أوجد σ^2 للمجتمع.

ب- أوجد تباين متوسط العينة $v(\bar{y})$.

ت- تحقق من العلاقة $v(\bar{y}) = \frac{\sigma^2}{n} \left(\frac{N-n}{N-1} \right)$.

2. لتقدير تركيز التيروكسين لدى المجتمع المؤلف من المواليد الذكور في مشفى التوليد العام الماضي والبالغ عدده $N = 1000$ مولود ، اخترنا عينة عشوائية بسيطة حجمها $n = 50$ فوجدنا أن $\bar{x} = 9.8$ والانحراف المعياري $S = 3.1$.

أ- احسب حد الخطأ في تقدير المتوسط.

ب- إذا أردنا ألا يتجاوز حد الخطأ القيمة 0.6 فكم يجب أن يكون حجم العينة مع اعتبار $\sigma = 3$ ؟

3. في إحدى المداجن تمت تربية 2000 صوص. لتقدير كمية (وزن) جميع الفرائج أخذنا عند البيع عينة من 50 فروجاً ، وحسبنا متوسط وزنها $\bar{x} = 3.2 \text{ k.g}$ والانحراف المعياري $S = 0.3$.

أ- قدر كمية الفروج الذي سيباع.

ب- احسب حد خطأ التقدير.

4. في اختبار لتعيين مدى فعالية دواء جديد لمرض معين أخذت عينة حجمها $n = 40$ من المصابين بهذا المرض الذين جربوا الدواء والبالغ عددهم $N =$

1000 فكان عدد الذين شفوا بتأثير هذا الدواء 30 مريضاً.

أ- كم تقدر النسبة الحقيقية P للذين تم شفاؤهم؟

ب- ما حد الخطأ لهذا التقدير؟

ت- ما حجم العينة اللازم أخذها كي لا يتجاوز حد الخطأ القيمة 0.04 ؟

5. شارك في برنامج رياضي وصحي لتخفيف الوزن 200 فتاة. اخترنا بشكل عشوائي 20 فتاة وحسبنا متوسط وانحراف الوزن المخفض للعينة فكانت $\bar{x} = 3.4 \text{ k.g}$ و $S = 1.2$.

أ- كم تقدر المتوسط الحقيقي μ للوزن المخفض للمجتمع؟

ب- احسب حد خطأ التقدير.

ت- أوجد حجم العينة اللازم سحبها كي لا يتجاوز حد الخطأ 0.5 .

6. أخذنا عينة عشوائية من 40 مريضاً من مجموعة المرضى الذين خضعوا للتجربة وعددهم 400 والذين يتناولون المضادات الحيوية بكثرة ، فوجدنا أن 10 منهم أصيب بتضخم عضلة القلب .

أ- قدر نسبة الذين سيصابون بتضخم عضلة القلب من الذين خضعوا للتجربة.

ب- احسب حد الخطأ.

ت- أوجد 95% مجال ثقة للنسبة P .